*European Journal of Endocrinology*, 2025, **193**, 188–196
https://doi.org/10.1093/ejendo/lvaf141
Advance access publication 14 July 2025
**Original Research**

# Impact of old and current immunoassays on the 1 mg overnight dexamethasone suppression test: comparison with LC-MS/MS

**Laura Rotolo,[1,2] Greta Galante,[2,3] Kimberly Coscia,[1,2] Valentina Bissi,[2,3] Lorenzo Tucci,[1,2] Marco Mezzullo,[2,3] Alessandra Gambineri,[1,2] Valentina Vicennati,[1,2] Guido Zavatta,[1,2] Uberto Pagotto,[1,2,3] Guido Di Dalmazi,[1,2,3] and Flaminia Fanelli[2,3,*]**

[1]Division of Endocrinology and Diabetes Prevention and Care, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna 40138, Italy
[2]Department of Medical and Surgical Sciences, Alma Mater Studiorum University of Bologna, Bologna 40138, Italy
[3]Center for Applied Biomedical Research, Department of Medical and Surgical Sciences, Alma Mater Studiorum University of Bologna, Bologna 40126, Italy

*Corresponding author: Department of Medical and Surgical Sciences, University of Bologna, S. Orsola Policlinic, Via Massarenti 9, Bologna 40138, Italy. Email: flaminia.fanelli2@unibo.it

## Abstract

**Objective:** The influence of immunoassay performance in hypercortisolism and dexamethasone suppression test (DST) settings was scarcely investigated. We evaluated the effectiveness of 2 immunoassays in detecting hypercortisolism compared to recommended liquid chromatography-tandem mass spectrometry (LC-MS/MS), and compared immunoassay analytical performance in basal and post-DST conditions.

**Methods:** We measured cortisol in post-DST sera of patients with suspected hypercortisolism or adrenal incidentalomas by Elecsys gen I ($n = 260$), and by Access ($n = 217$). All samples were also measured by a validated LC-MS/MS method. We estimated hypercortisolism rate according to the established 50 nmol/L cutoff, and generated immunoassay-specific cutoffs providing >95% sensitivity and >80% specificity. Finally, we compared cortisol measurements in basal and post-DST samples.

**Results:** Using the 50 nmol/L cutoff, both immunoassays detected lower rates of hypercortisolism compared with LC-MS/MS, particularly in patients with adrenal adenomas ($P < .050$). Elecsys gen I and Access determined 6.9% and 6.4% possible false negatives, respectively. Elecsys gen I also caused 3.8% possible false positives. Optimal cutoff was 41 nmol/L for Elecsys gen I (sensitivity: 97.7%; specificity: 80.8%), and 33 nmol/L for Access (sensitivity: 97.5%; specificity: 78.3%). In basal and post-DST samples, Elecsys gen I overestimated by 32.5% and 6.1%, whereas Access underestimated by −4.7% and −5.9% compared to LC-MS/MS cortisol measurements, respectively. Sex differences in method deviations were noted.

**Conclusions:** Both immunoassays demonstrated remarkable underdetection of hypercortisolism, suggesting the application of a method-specific cutoff. Immunoassay performance may not be uniform in basal and post-DST conditions and should be purposely examined. Accurate LC-MS/MS methods should be preferred in hypercortisolism settings.

**Keywords:** dexamethasone suppression test, hypercortisolism, adrenal incidentalomas, immunoassay, liquid chromatography—tandem mass spectrometry

### Significance

The impact of immunoassay performance on cortisol measurement was evaluated in multiple settings. However, scarce attention was paid to hypercortisolism and dexamethasone suppression test (DST). We compared 2 routine immunoassays with a validated LC-MS/MS method in serum from patients with suspected hypercortisolism or adrenal incidentalomas. We found that immunoassays performed differently in basal and post-DST samples, and were differently affected by cross-reactivity, insufficient corticosteroid-binding globulin displacement and random error. By using the post-DST cortisol cutoff of 50 nmol/L suggested by guidelines, both immunoassays exhibited low sensitivity in detecting hypercortisolism, particularly in adrenal benign masses. Method-specific cutoff could improve sensitivity. However, accurate LC-MS/MS is preferable in DST setting, as it offers elevated specificity and promises the final standardization of cutoff.

**Received:** January 25, 2025. **Revised:** June 22, 2025. **Editorial Decision:** July 7, 2025. **Accepted:** July 10, 2025

## Introduction

The overnight 1 mg dexamethasone suppression test (DST) is used as a first-line test for the assessment of hypercortisolism.[1-3] The synthetic glucocorticoid acts on brain structures to elicit the self-feedback regulation of the hypothalamus–pituitary–adrenal (HPA) axis, thereby ultimately lowering cortisol production in the absence of hypercortisolism. Cortisol levels of 50 nmol/L or lower the morning after dexamethasone intake indicate an effective capacity of HPA self-regulation. Conversely, cortisol levels above the cutoff suggest the condition of hypercortisolism.[1-5] The 50 nmol/L cutoff was determined by Blethen and Chasalow in 1989 by using a radioimmunoassay (RIA).[5] Soon after, automated immunoassays, prone to several drawbacks, evolved. More recently, liquid chromatography-tandem mass spectrometry (LC-MS/MS) methods were introduced. According to guidelines, such cutoff provides sensitivity and specificity above 95% and 80%, respectively, for hypercortisolism detection.[1] However, the comparability of current methods with the original RIA is unknown, posing the patients at possible risk of under- or over-investigation.

Cortisol immunoassays have repeatedly shown suboptimal reproducibility and poor accuracy caused by antibody cross-reactivity with cortisol metabolites and lack of standardization.[6-8] Moreover, suboptimal analyte displacement from corticosteroid-binding globulin (CBG) was often found to determine underestimation of measurements in a sex-dependent fashion.[6-8] At variance, a few studies reported good accuracy and interlaboratory comparability of LC-MS/MS methods.[7,9] However, these studies focused on physiological or supraphysiological cortisol levels, as achieved with Synacthen® stimulation.[6-9] Furthermore, materials distributed in external quality control surveys are most often characterized by cortisol in middle or high circulating range. Unfortunately, method performance may vary at low levels elicited by the DST, and it is therefore necessary to obtain information at the exact levels at which clinical decisions about hypercortisolism are made. At present, there is scarce information about performance of immunoassay in the DST cortisol range, and about how outcomes of patients screened for hypercortisolism because of different pathological conditions may change depending on the assay used.[10]

Our study aimed to utilize a validated LC-MS/MS method to estimate the efficacy of 2 immunoassays in detecting hypercortisolism at the DST, and their analytical performance in hypercortisolism setting.

## Methods

### Study cohort

This is a single-center retrospective study. The local ethics committee approved the study (protocol 45/2016/O/Oss). We enrolled 477 adult outpatients referring to the Division of Endocrinology and Diabetes Prevention and Care of the S. Orsola Polyclinic of Bologna for clinical suspicion of hypercortisolism or for the hormonal evaluation of adrenal incidentalomas. Pregnant or breastfeeding women and patients affected by conditions such as gastrointestinal malabsorption, nephrotic syndrome, liver failure, or taking medications influencing dexamethasone or cortisol metabolism (eg, exogenous glucocorticoids, estro-progestogens, CYP3A4 inductors or inhibitors) within 3 months before the evaluation were excluded. All gave written informed consent for participation. All the procedures conducted during this study were performed in compliance with the Declaration of Helsinki.

### Clinical evaluations

Body mass index, waist circumference, and blood pressure were registered, along with fasting glycolipid profile, creatinine and transaminases. The presence of hypertension, diabetes, low bone mineral density, or osteoporosis was deduced from medical records. All patients underwent an overnight 1 mg DST as per clinical practice (Decadron® 0.5 mg, Caber Farmaceutici, Pomezia, Italy; 2 tablets taken at 23.00-24.00 PM).[1-3] The 24-hour urinary free cortisol (UFC) and late-night salivary cortisol (LNSC) were performed in patients with suspected clinical hypercortisolism. If confirmed, adrenocorticotropic hormone (ACTH) measurement and radiological investigations were used to differentiate between adrenal or pituitary hypersecretion.[1,2]

### Laboratory methods

All patients provided blood both in basal conditions and the morning after (8.00-9.00 AM) the dexamethasone intake. Fresh serum samples were measured by the routine immunoassay available at the time of sample collection. Specifically, samples collected between 2013 and 2014 ($n = 260$) were measured by the Cortisol Elecsys test gen I on the Modular Analytics E170 (Roche Diagnostics GmbH, Mannheim, Germany) at the Laboratorio Centralizzato of the S.Orsola Hospital of Bologna, whereas samples between 2022 and 2023 ($n = 217$) were measured by the Access Cortisol on the DxI 800 (Beckman Coulter, Brea, California), at the Laboratorio Unico Metropolitano, Azienda Sanitaria Locale of Bologna. In 2013 and 2022, we received funding for research studies requiring LC-MS/MS measurements for other purposes. Therefore, we stored a second aliquot at −80 °C. LC-MS/MS measurement was performed within the following 6 months from serum storage.

The Cortisol Elecsys gen I is a competitive electrochemiluminescence immunoassay using a polyclonal antibody and ruthenium-labelled cortisol as competitor. Danazol is used for CBG displacement. Traceability was obtained against the Enzymun-test cortisol, which was standardized against an isotopic dilution MS method. Accuracy was 89%-111% in materials from the IFCC-451 panel of the reference materials and measurements (cortisol levels between 83 and 764 nmol/L). Measurement range was 0.5-1750 nmol/L, functional sensitivity was 8.5 nmol/L, and intra-assay and interassay CV range were below 1.7% and 2.8%, respectively (range 124-866 nmol/L). Cross-reactivity was reported from 6-a-methylprednisolone (389%), prednisolone (171%), and allotetrahydrocortisol (165%), tested at 270 nmol/L; 6-β-hydroxycortisol (158%) and 21-deoxycortisol (45.4%) at 2700 nmol/L; and corticosterone (5.8%), 11-deoxycortisol (4.1%), and 17OH-progesterone (1.5%) at 27 000 nmol/L.

The Access cortisol is a chemiluminescent competitive immunoenzymatic assay (CLIA) using a polyclonal anticortisol rabbit antibody, paramagnetic beads coated with goat anti-rabbit polyclonal antibody, and alkaline phosphatase-conjugated cortisol as competitor. Measurement range was 11-1655 nmol/L, and functional sensitivity was 11 nmol/L. Accuracy was 97.2%-109.0% as determined by standard addition (36-1606 nmol/L range). The intra-assay and total CVs

were below 6.7% and 7.9%, respectively (166-1059 nmol/L range). Cross-reactivity was reported from prednisolone (23.9%) tested at 555 nmol/L; 11-deoxycortisol (17.8%), cortisone (8.1%), and corticosterone (2.1%) at 2800 nmol/L; and 17OH-progesterone (5.3%) and prednisone (3.1%) at 29 000 nmol/L. Traceability and CBG displacement were not reported.

The LC-MS/MS method was previously validated[11] at the Center for Applied Biomedical Research, University of Bologna, Italy, and recently verified within the HarmoSter survey.[9] Briefly, 600 µL of serum were purified by solid phase extraction. Extracts were separated on a Luna RP-C8 100 × 4.6 mm, 5 mm (Phenomenex, Torrance, California) in a 21-min gradient run on the Serie 200 HPLC (PerkinElmer, Waltham, Massachusetts). MS-detection was obtained on the API 4000-QTrap (Sciex, Framingham, Massachusetts) by using 363.2/267.0 and 367.3/97.1 transitions for cortisol and d4-cortisol, respectively. Isotopic dilution and ion ratio monitoring ensured accurate quantification. Measurement range was 0.66-1379.5 nmol/L. Performance was verified on certified quality control materials provided by the Reference Institute for Bioanalytics (Bonn, Germany). Mean bias and CV calculated across a 10-year study period were 1.3% and 3.3% in the 262-284 nmol/L range and 1.8% and 3.8% in the 524-731 nmol/L range, respectively. Other adrenal metabolites measured by LC-MS/MS were cortisone, 11-deoxycortisol, corticosterone, 17OH-progesterone and progesterone.[11]

ACTH was measured by CLIA (Immulite 2000; Siemens Healthcare Diagnostics Inc, Munich, Germany, reference interval: 5-60 pg/mL). The 24-hour UFC was measured with Elecsys gen I until 2014 and with Access since 2015. LNSC was measured by the Biological Sales Network kit for LC-MS/MS (Castelleone, Italy). Creatinine, total cholesterol, triglycerides, glucose, and insulin were measured by Roche Cobas 8000 or Elecsys Modular Analytics for cohort 1 and by Beckman Coulter AU5800 or DxI analyzers for cohort 2. Glycated hemoglobin was measured by HPLC (Bio-Rad, Hercules, CA).

### Study design and statistics

Each patient provided paired basal and post-DST samples. Each sample was measured by both routine immunoassay and LC-MS/MS. Cohort 1 included patients recruited between 2013 and 2014 who had their routine cortisol measured by Elecsys gen I ($n = 260$). Cohort 2 included patients recruited between 2022 and 2023 who had their routine cortisol measured by Access ($n = 217$). Patients were further subgrouped according to clinical and radiological findings leading to the diagnosis of adrenal adenoma, adrenal hyperplasia, adrenocortical carcinoma, other adrenal masses (myelolipoma, adrenal metastases, lymphoma, pheochromocytoma, adrenal cyst, and schwannoma), or Cushing's disease. A further group included patients with no known adrenal or pituitary tumor in whom the suspicion of Cushing's syndrome was afterward clinically or biochemically excluded, and patients formerly affected by Cushing's syndrome or disease in complete remission after successful surgery (Table 1).

Considering LC-MS/MS as the reference and 50 nmol/L as the cutoff,[1-3] we calculated possible false positives and false negatives provided by the 2 immunoassays, within the 2 overall cohorts and each clinical subgroup. Within-subject (CVi) and between-subjects (CVg) biological variability available for basal cortisol levels[12] were used to calculate total allowable error ($TAE = 0.25 \times (CVi^2 + CVg^2)^{0.5} + 1.65 \times (0.5 \times CVi)$) at 22.2%.[13] Biological variability for post-DST cortisol is not known.[12] Method comparison was performed between Elecsys gen I and LC-MS/MS in cohort 1, and between Access and LC-MS/MS in cohort 2. Data were analyzed by Mann–Whitney, Wilcoxon, Bland–Altman, McNemar, and ROC tests, Spearman's rank correlation and Passing–Bablok regression. Cross-reactivity from metabolites was evaluated by multiple regression analysis including cortisol, cortisone, 11-deoxycortisol, corticosterone, 17OH-progesterone, and progesterone as independent variables, and the difference between immunoassay and LC-MS/MS as dependent variable. Analyses were performed by MedCalc® Statistical Software (v.20.104, Ostend, Belgium). $P \leq .050$ was considered statistically significant.

## Results

Demographic, diagnostic, and clinical characteristics of the 2 cohorts are reported in Table 1 and Table S1 and S2. Cases in each cohort were subgrouped in nonsecreting (post-DST cortisol ≤ 50 nmol/L) or hypercortisolism (post-DST cortisol > 50 nmol/L) as determined either with immunoassay or LC-MS/MS. In cohort 1, the prevalence of hypercortisolism was similar when measured by Elecsys gen I or LC-MS/MS (Table 2). However, when considering adrenal adenomas, the prevalence of hypercortisolism with Elecsys gen I was significantly lower than LC-MS/MS (26.8% vs 34.8%; $P = .049$). Cases with discordant hormonal assessment were 10.8% ($n = 28$ of 260), including 3.8% ($n = 10$) possible false positives and 6.9% ($n = 18$) possible false negatives (Table 2; Figure 1A and C). Among possible false positives, % differences between Elecsys gen I and LC-MS/MS ranged between 26.9% and 156.9%, all above the TAE of 22.2%. Among possible false negatives, % differences were between −35.6% and −2.7%, of which 27.8% ($n = 5$ of 18) showed a bias larger than the TAE of −22.2% (Figure S1A and 1B). Notably, possible false positives within other adrenal masses and Cushing's syndrome excluded groups accounted for 5.9% ($n = 1$ of 17) and 9.5% ($n = 2$ of 19), respectively, whereas possible false negatives within adrenal adenomas were 11.6% ($n = 13$ of 112) (Table 2).

In cohort 2 (Table 3), the overall prevalence of hypercortisolism was lower by Access compared with LC-MS/MS (30.4% vs 36.4%, respectively; $P = .001$), particularly within adrenal adenomas (30.9% vs 40.4%, respectively; $P = .004$). Discordant cases accounted for 6.9% ($n = 15$ of 217), including 0.5% ($n = 1$) possible false positives and 6.4% ($n = 14$) possible false negatives (Table 3; Figure 1B and D). The % difference between Access and LC-MS/MS in the only possible false positive case was 77.0%, whereas it ranged −59.1%–−2.5% among possible false negatives, with 64.3% ($n = 9$ of 14) of these showing a bias larger than the TAE of −22.2% (Figure S1C and 1D). Remarkably, possible false negatives represented 9.6% ($n = 9$ of 94) of adrenal adenomas and 6.8% ($n = 5$ of 73) of adrenal hyperplasias (Table 3).

As reported in Table S3, ROC analysis revealed an AUC of 0.958 for Elecsys gen I and 0.970 for Access. The cutoff set at 50 nmol/L determined sensitivity and specificity of 79.6% and 94.2% for Elecsys gen I in cohort 1, and of 82.3% and 99.3% for Access in cohort 2, respectively. Considering the suggested

**Table 1.** Demographic and diagnostic classification of the 2 study cohorts.

| | Cohort 1 | | Cohort 2 | |
|---|---|---|---|---|
| | *N* (%) | Age (years; median (IQR)) | *N* (%) | Age (years; median (IQR)) |
| All | 260 (100) | 62.9 (51.0-70.1) | 217 (100) | 62.4 (50.1-69.9) |
| Adrenal adenoma | 112 (43.1) | 62.3 (51.7-69.2) | 94 (43.3) | 63.8 (54.1-69.7) |
| Adrenal hyperplasia | 106 (40.8) | 62.3 (55.1-72.6) | 73 (33.6) | 64.8 (54.6-71.1) |
| Adrenocortical carcinoma | 2 (0.8) | 56.9 (56.4-57.4) | 5 (2.3) | 50.0 (40.7-55.8) |
| Cushing's disease | 2 (0.8) | 47.9 (46.7-49.2) | 3 (1.4) | 48.2 (32.9-53.7) |
| Other adrenal mass | 17 (6.5) | 65.7 (57.2-70.4) | 13 (6.0) | 59.1 (56.0-66.4) |
| Cushing's syndrome excluded | 21 (8.0) | 44.8 (42.3-51.8) | 29 (13.4) | 44.0 (36.1-53.1) |

**Table 2.** Classification of the functional status of patients in cohort 1 based on post-DST cortisol measured by Elecsys gen I and LC-MS/MS, according to the 50 nmol/L cutoff.

| Study groups | Elecsys gen I | | LC-MS/MS | | *P*-value[a] | Concordant cases (*N*; %) | Discordant cases | | |
|---|---|---|---|---|---|---|---|---|---|
| | NS (*N*; %) | HC (*N*; %) | NS (*N*; %) | HC (*N*; %) | | | Total (*N*; %) | Possible false positives (*N*; %) | Possible false negatives (*N*; %) |
| All ($N = 260$) | 180; 69.2 | 80; 30.8 | 172; 66.2 | 88; 33.8 | 0.185 | 232; 88.9 | 28; 10.8 | 10; 3.8 | 18; 6.9 |
| Adrenal adenoma ($N = 112$) | 82; 73.2 | 30; 26.8 | 73; 65.2 | 39; 34.8 | 0.049 | 95; 84.8 | 17; 15.2 | 4; 3.6 | 13; 11.6 |
| Adrenal hyperplasia ($N = 106$) | 65; 61.3 | 41; 38.7 | 63; 59.4 | 43; 40.6 | 0.727 | 98; 92.5 | 8; 7.5 | 3; 2.8 | 5; 4.7 |
| Adrenocortical carcinoma ($N = 2$) | 1; 50 | 1; 50 | 1; 50 | 1; 50 | n.a. | 2; 100 | 0; 0 | 0; 0 | 0; 0 |
| Cushing's disease ($N = 2$) | 0; 0 | 2; 100 | 0; 0 | 2; 100 | n.a. | 2; 100 | 0; 0 | 0; 0 | 0; 0 |
| Other adrenal mass ($N = 17$) | 13; 76.5 | 4; 23.5 | 14; 82.4 | 3; 17.6 | 1.000 | 16; 94.1 | 1; 5.9 | 1; 5.9 | 0 |
| Cushing's syndrome excluded ($N = 21$) | 19; 90.5 | 2; 9.5 | 21; 100.0 | 0; 0 | 0.500 | 19; 90.5 | 2; 9.5 | 2; 9.5 | 0 |

Percentages refer to the total number of cases in each clinical subgroup.
[a]McNemar test.
NS, nonsecreting; HC, hypercortisolism; n.a., not available.

sensitivity and specificity above 95% and 80%, respectively, for hypercortisolism screening,[1] the suitable cutoff for Elecsys gen I was 41 nmol/L (97.7% sensitivity and 80.8% specificity) (Table S3). For Access, ideal sensitivity was found with cutoff at 33 nmol/L (97.5% sensitivity and 78.3% specificity), and ideal specificity at 39 nmol/L (92.4% sensitivity and 89.1% specificity) (Table S3). No relevant differences were noted in males and females.

In cohort 1, a large overestimation of 32.5% (95% CI: 30.4%-34.7%) was observed for Elecsys gen I compared with LC-MS/MS in basal samples (Figure S2A and Table S4), with median difference of 95.46 (64.70-143.33) nmol/L ($P < .001$, Table S5). In post-DST samples, overestimation reduced to 6.1% (2.8%-9.4%) (Figure S2B and Table S4). A sex effect was noted, causing larger overestimation in males than in females both in basal ($P = .039$) and post-DST samples ($P < .001$) (Table S5). In particular, the difference between Elecsys gen I and LC-MS/MS in post-DST samples was 14.9% (9.0%-20.7%) in males, with a median difference of 2.65 (−0.46-7.66) nmol/L ($P < .001$; Table S4 and S5), and not significant in females. In cohort 2, Access bias vs LC-MS/MS was −4.7% (−7.3%−−2.0%) with a median difference of −8.83 (−54.73-19.56) nmol/L ($P < .001$) in basal, and −5.9% (−9.5%−−2.4%) with a median difference of −3.92 (−10.50-1.50) nmol/L ($P < .001$) in post-DST samples (Figure S2C and S2D; Table S4 and S5). Post-DST, underestimation differed by sex ($P = .032$), and was −7.6% (−12.1%−−3.2%) in females, with a median difference of −5.82 (−10.94-0.70) nmol/L ($P < .001$), and not significant in males (Tables S4 and S5).

In cohort 1, the comparison of Elecsys gen I with LC-MS/MS in basal samples resulted in rank correlation coefficient of 0.935 (0.917-0.948), large proportional overestimation (slope coefficient (95% CI): 1.441 (1.385-1.504)), and large constant underestimation (intercept coefficient (95% CI): −36.05 (−54.25–−19.15)) (Figure 2A and Table S6). In post-DST samples, correlation coefficient was 0.905 (0.881-0.925) and a small proportional overestimation was noted (slope (95% CI): 1.055 (1.000-1.115)) (Figure 2B and Table S6). Comparing Access with LC-MS/MS in cohort 2, correlation coefficient of 0.845 (0.802-0.870) was found in basal samples, with no significant bias (Figure 2C and Table S6). However, a proportional underestimation was found in post-DST samples (slope (95% CI): 0.885 (0.838-0.935)), with $r = 0.893$ (0.862-0.917) (Figure 2D and Table S6). No relevant differences were noted elaborating data within clinical subgroups (Table S6).

In cohort 1, steroids directly associated with Elecsys gen I overestimation over LC-MS/MS in basal samples were cortisone ($t = 2.40$, $P = .019$), 11-deoxycortisol ($t = 4.74$, $P < .001$), and corticosterone ($t = 2.12$, $P = .037$). In cohort 2, Access bias vs LC-MS/MS in basal samples was inversely associated with 17OH-progesterone ($t = -5.47$, $P < .001$), and positively associated with corticosterone ($t = 5.66$, $P < .001$). No associations were found in post-DST samples of both cohorts.

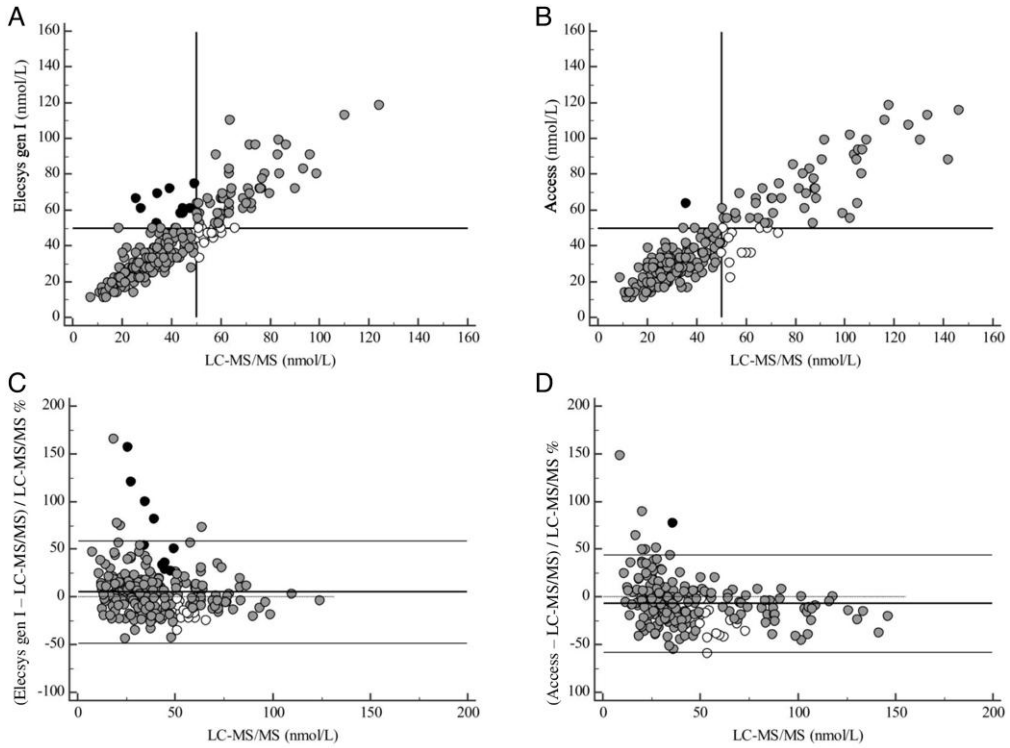


**Figure 1.** Close-up view of agreement between immunoassays and LC-MS/MS in post-DST cortisol measurements from cohort 1 and 2. For clarity reasons, only samples with post-DST cortisol <120 nmol/L are shown. (A and C) Comparison between Elecsys gen I and LC-MS/MS in cohort 1; (B and D) comparison between Access and LC-MS/MS in cohort 2. (A and B) Close-up view of the regression analysis. Black lines: cutoff at 50 nmol/L; (C and D) close-up view of the Bland–Altman analysis. Bold and thin black lines: mean difference and 95% agreement interval, respectively; dotted line: line of best fit, 0%. Gray circles: samples in which immunoassay and LC-MS/MS gave concordant results as to the presence or absence of hypercortisolism; black circles: possible false positives by immunoassays; white circles: possible false negatives by immunoassays.

## Discussion

Several studies investigated the accuracy of serum cortisol measurements by various immunoassays in different settings, including pregnancy, estro-progestogen users, kidney diseases, septic shock, intensive care, adrenal insufficiency, and metyrapone treatment.[7,8,14-21] Notably, the application of method-specific cortisol cutoffs for basal and Synacthen testing was repeatedly suggested in the context of adrenal insufficiency.[16-18,20,21] However, performance issues described in basal and post-Synacthen samples cannot be compared to DST because of differences in analyte levels, in the milieu of cross-reacting metabolites, and in the proportion between total cortisol and CBG.

The only study addressing the influence of immunoassay performance on DST was recently published by Atkins et al.[10] Notably, they investigated 3 popular immunoassays. Our study now adds information about 2 other widespread methods, thereby increasing awareness among clinicians about the heterogeneity determined by different immunoassays in the context of hypercortisolism management. Elecsys gen I was on the market from 2000 to 2015, whereas Access has been in use for about fifteen years. Both assays were used in several laboratories worldwide and in several studies on Cushing's syndrome and adrenal incidentalomas.[6,22-25]

By using the recommended 50 nmol/L cutoff, both immunoassays exhibited a sensitivity around 80%, which is relevantly lower than the 95% suggested for the screening of hypercortisolism.[1-4] Indeed, compared with LC-MS/MS, both Elecsys gen I and Access could overall result in up to 7% missed cases of hypercortisolism, which could rise to 10%-12% among patients with benign cortical adrenal masses. Of note, the risk of under-detection of hypercortisolism appeared higher for Access than Elecsys gen I, as most of the possible false negatives of the former displayed a bias larger than the TAE. Considering an average of 10 patients performing the 1 mg overnight DST every week, and that adrenal incidentalomas represent the vast majority of patients undergoing this test, Access is possibly causing up to 40 patients under-investigated every year in our unit, all potentially lost at follow-up. Further concerns arise when considering that possible false positives occurred in 10% of patients in which Cushing's syndrome was afterward excluded according to the overall clinical, radiological, and laboratory evaluations, and in up to 6% of patients affected by other types of adrenal masses. The recommended sensitivity above 95% for hypercortisolism screening could be achieved by setting the cutoff at 41 nmol/L for Elecsys gen I and at 33 nmol/L for Access. The latter could improve the management of hypercortisolism in centers currently using the Access method. The former still represents useful information for retrospective evaluations of patients who were evaluated with the old Elecsys assay. On the other hand, the specificity obtained with these cutoffs appeared suboptimal, indicating the need for more reliable assays.

The 3 immunoassays tested by Atkins et al. in post-DST samples performed very differently. Unfortunately, the limited number of cases and the lack of clinical information in their study did not allow to define categories of patients particularly exposed to consequences of poor method performance.[10]

**Table 3.** Classification of the functional status of patients in cohort 2 based on post-DST cortisol measured by Access and by LC-MS/MS, according to the 50 nmol/L cutoff.

| Study groups | Access | | LC-MS/MS | | *P*-value[a] | Concordant cases | Discordant cases | | |
|---|---|---|---|---|---|---|---|---|---|
| DD | NS (*N*; %) | HC (*N*; %) | NS (*N*; %) | HC (*N*; %) | | Total (*N*; %) | Total (*N*; %) | Possible false positives (*N*; %) | Possible false negatives (*N*; %) |
| All (*N* = 217) | 151; 69.6 | 66; 30.4 | 138; 63.6 | 79; 36.4 | 0.001 | 202; 93.1 | 15; 6.9 | 1; 0.5 | 14; 6.4 |
| Adrenal adenoma (*N* = 94) | 65; 69.1 | 29; 30.9 | 56; 59.6 | 38; 40.4 | 0.004 | 85; 90.4 | 9; 9.6 | 0; 0 | 9; 9.6 |
| Adrenal hyperplasia (*N* = 73) | 43; 58.9 | 30; 41.1 | 39; 53.4 | 34; 46.6 | 0.219 | 67; 93.2 | 6; 8.2 | 1; 1.4 | 5; 6.8 |
| Adrenocortical carcinoma (*N* = 5) | 1; 20.0 | 4; 80.0 | 1; 20.0 | 4; 80.0 | n.a. | 5; 100.0 | 0; 0 | 0; 0 | 0; 0 |
| Cushing's disease (*N* = 3) | 0; 0 | 3; 100 | 0; 0 | 3; 100 | n.a. | 3; 100.0 | 0; 0 | 0; 0 | 0; 0 |
| Other adrenal mass (*N* = 13) | 13; 100.0 | 0 | 13; 100.0 | 0 | 1.000 | 13; 100.0 | 0; 0 | 0; 0 | 0; 0 |
| Cushing's syndrome excluded (*N* = 29) | 29; 100.0 | 0 | 29; 100.0 | 0 | 1.000 | 29; 100.0 | 0; 0 | 0; 0 | 0; 0 |

Percentages refer to the total number of cases in each clinical subgroup.
[a]McNemar test.
NS, nonsecreting; HC, hypercortisolism; n.a., not available.

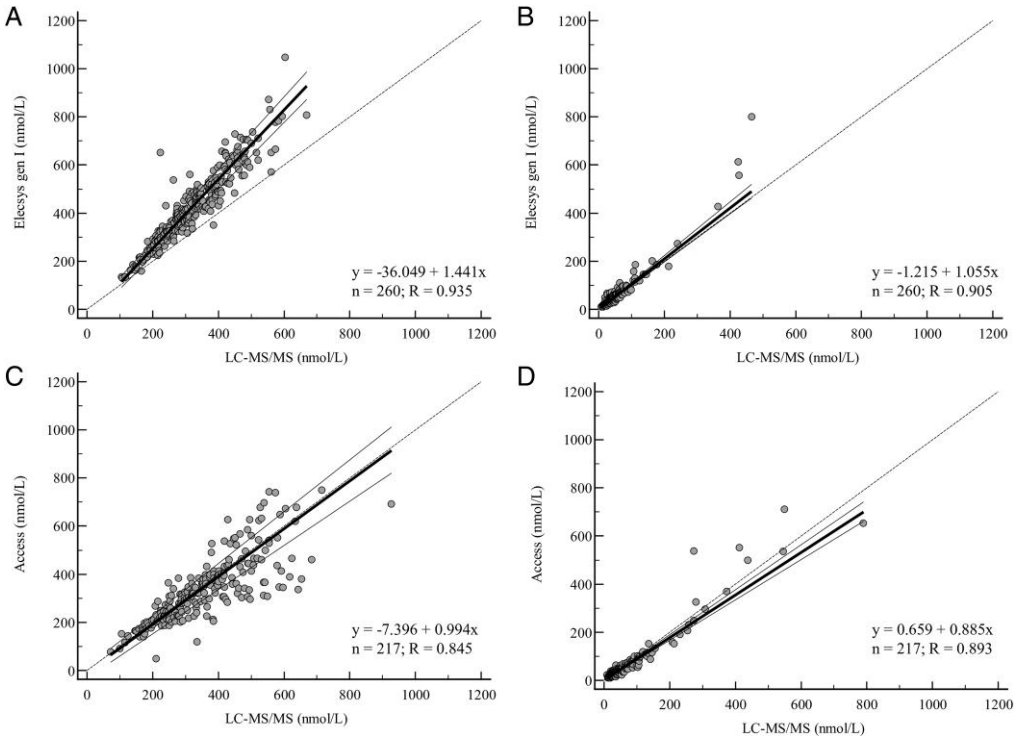


**Figure 2.** Passing and Bablok regression for cortisol measurement by immunoassays and LC-MS/MS. Regression in basal (A and C) and post-DST samples (B and D) are reported for comparison of Elecsys gen I vs LC-MS/MS in cohort 1 (A and B), and of Access vs LC-MS/MS in cohort 2 (C and D) vs LC-MS/MS. Bold and thin lines represent regression lines and 95% CI, respectively. Dashed line represents the line of best fit.

The Elecsys gen II showed the best agreement with LC-MS/MS,[10] thereby supporting previous evidence about significant ameliorations introduced in the current Elecsys assay.[7,17,18,20,21,26] At variance, Abbott Alinity and Siemens ADVIA Centaur caused 19.2% possible false negatives and 7.5% possible false positives, respectively. In line with our findings, a lower cutoff of 31.2 nmol/L was suggested for Alinity.[10] Of note, 3 out of the 5 immunoassays tested by ourselves and Atkins required a non-negligible reduction of the cutoff.

Our findings supporting the lowering of Elecsys gen I cutoff seemed in contrast with previous literature reporting a large overestimation of cortisol levels, between 30% and 50%, attributed to miscalibration and cross-reactivity from steroidal species.[17-21,26] One potential explanation for this discrepancy lies in the fact that previous studies focused on basal samples, whereas our cutoff was defined in suppressed conditions.[17-21,26] To resolve this discrepancy, we compared immunoassays and LC-MS/MS measurements separately in basal and post-DST paired samples. While in basal samples our findings confirmed the large overestimation reported in the literature,[17-21,26] in post-DST samples Elecsys gen I only reported a modest bias in comparison with our LC-MS/MS method. Overestimation might be mitigated in post-DST samples due to the incomplete displacement of cortisol from CBG,

whose negative impact at suppressed cortisol levels is expectedly higher. Indeed, while most of the CBG capacity is saturated in basal conditions, leaving a greater proportion of free cortisol accessible to assay antibody, after DST a larger proportion of hormone is bound to the transport protein.[8] In support of this hypothesis, and in line with CBG levels being slightly higher in women than in men,[27] we noted that Elecsys gen I overestimation is higher in males than females. Sex effect seemed even amplified in post-DST samples, where average overestimation in males accounted for 15%, but was not significant in females. Another reason for the bimodal performance of Elecsys gen I could derive from the suppression of cross-reacting metabolites upon dexamethasone administration. Supporting this, we found a systematic contribution from cortisone, 11-deoxycortisol and corticosterone to Elecsys gen I overestimation in basal but not in post-DST conditions. We could speculate that cross-reactivity may explain part of the large TAE observed in the sizable number of possible false positive cases of hypercortisolism detection.

Overall, our data on Access were in line with data previously reported in literature.[18,19] Access exhibited a suboptimal correlation with LC-MS/MS, suggesting the presence of random errors from the former, which can only minorly be explained by cross-reactivity toward steroid metabolites. Moreover, a slight underestimation was noted in post-DST samples, which, upon deeper analysis, was only significant in females. In line with available data,[7,16] the CBG displacement issue in Access appeared less severe than in Elecsys gen I. In both cases, however, sex did not impact the effectiveness of hypercortisolism detection.

Unfortunately, this study relied on measurements performed in singlet. Therefore, we could not dissect the independent contribution of assays imprecision from sources of bias, including calibration, cross-reactivity from steroid metabolites, CBG displacement or others, in individual measurements and, in particular, in cases of misclassification. On the other hand, our study benefited from the large number of patients and detailed clinical classification. This allowed us to reliably estimate the rate of possible misdetection, to identify categories of patients particularly interested by this issue, and to provide method-specific cutoffs. However, the 2 immunoassays were compared with the same LC-MS/MS method in 2 independent cohorts. Though this represents a potential confounding element, we believe that the wide number of cases and the similar distribution within clinical subgroups prevented any relevant artifact.

Unfortunately, we did not have the possibility to measure dexamethasone serum concentration in post-DST samples to verify drug intake and absorption and therefore the reliability of the DST outcome. Nonetheless, the verification of dexamethasone levels would not have changed the comparability of cortisol measurement between immunoassays and LC-MS/MS.

Another limitation of our study, as in previous ones in literature, consists in the lack of an independent test to determine hypercortisolism. Therefore, even though our LC-MS/MS method was continuously verified through certified quality controls and demonstrated optimal precision, accuracy and comparability to other LC-MS/MS methods,[9] we cannot assess whether it guarantees optimal screening effectiveness at 50 nmol/L. Studies establishing reference limits of suppressed cortisol levels in large numbers of healthy subjects are required. Of note, a recent study by Vogg et al., focusing on patients with Cushing's syndrome and patients in which Cushing's syndrome was excluded, established a LC-MS/MS-specific post-DST cortisol cutoff of 66 nmol/L, providing 100% sensitivity and 92.4% specificity.[28] Standardization of cortisol measurement is mandatory before generalizing this new LC-MS/MS-specific cutoff among other LC-MS/MS users. Even though standardization appears to be an affordable goal,[9] several issues are limiting the path. First, no data are yet available about the reproducibility of LC-MS/MS methods for cortisol measurement in post-DST settings. Secondly, there is a lack of certified quality controls for suppressed cortisol levels to help the verification of method accuracy. Thirdly, there is no information about intraindividual and interindividual biological variability of suppressed cortisol levels to support the establishment of definitive performance requirements in terms of allowable imprecision, bias and total error, for LC-MS/MS methods measuring post-DST cortisol.[12] The scientific community should put effort in the achievement of such pillars as they will definitively improve the management of hypercortisolism.

In conclusion, the performance of immunoassays is not uniform across large analyte intervals and should be adequately investigated around particular clinical decision levels. In this sense, the performance of the old Elecsys gen I was better than expected at cortisol levels typical of DST. We highlighted that both Elecsys gen I and Access may fail in detecting hypercortisolism when the recommended 50 nmol/L cutoff is applied. Adjusting the cutoff in a method-specific manner could mitigate the risk of under-detection of hypercortisolism. However, immunoassays remain susceptible to interferences that hamper DST specificity and make the application of guideline-suggested cutoff impossible. Novel LC-MS/MS methods should be preferred in such critical specialist settings. Efforts are still needed to standardize LC-MS/MS measurements with the common aim of establishing a novel, universally applicable post-DST cortisol cutoff.

## Supplementary material

Supplementary material is available at *European Journal of Endocrinology* online.

## Funding

The study was supported by the Emilia-Romagna Region, Alessandro Liberati Young Researcher Grants (PRUA 1–2012–004), Italian Ministry of University and Research, Progetti di Rilevante Interesse Nazionale (PRIN) (2020ERLB52).

## Authors' contributions

Laura Rotolo (Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Software [equal], Validation [equal], Writing—original draft [equal], Writing—review & editing [equal]), Greta Galante (Data curation [equal], Formal analysis [equal], Writing—review & editing [equal]), Kimberly Coscia (Data curation [equal], Investigation [equal], Writing—review & editing [equal]), Valentina Bissi (Data curation [equal], Formal analysis [equal], Writing—review & editing [equal]), Lorenzo Tucci (Data curation [equal], Investigation [equal], Writing—review & editing [equal]), Marco Mezzullo (Data curation [equal], Formal analysis [equal], Writing—review & editing [equal]), Alessandra Gambineri (Investigation [equal], Writing—review

& editing [equal]), Valentina Vicennati (Investigation [equal], Writing—review & editing [equal]), Guido Zavatta (Data curation [equal], Investigation [equal], Writing—review & editing [equal]), Uberto Pagotto (Conceptualization [equal], Funding acquisition [equal], Methodology [equal], Writing—review & editing [equal]), Guido Di Dalmazi (Conceptualization [equal], Investigation [equal], Writing—review & editing [equal]), and Flaminia Fanelli (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Software [equal], Validation [equal], Writing—original draft [equal], Writing—review & editing [equal])

L.R. and F.F. performed the statistical analysis and wrote the manuscript. G.G., V.B., M.M., and F.F. measured the study samples. L.R., K.C., L.T., A.G., V.V., G.Z., and G.D.D. performed patient recruitment and examination. U.P., G.D.D., and F.F. conceived and designed the study. All contributed to manuscript finalization.

*Conflict of interest:* F.F. was appointed as a technical and scientific consultant by Roche Diagnostics International Ltd in 2023. All other authors report no conflict of interest in this work. A.G. is on the editorial board of EJE. She was not involved in the review or editorial process for this paper, on which she is listed as an author.

## Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

## References


1. Nieman LK, Biller BMK, Findling JW, *et al.* The diagnosis of Cushing's syndrome: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab*. 2008;93(5):1526-1540. https://doi.org/10.1210/jc.2008-0125
2. Fleseriu M, Auchus R, Bancos I, *et al.* Consensus on diagnosis and management of Cushing's disease: a guideline update. *Lancet Diabetes Endocrinol*. 2021;9(12):847-875. https://doi.org/10.1016/S2213-8587(21)00235-7
3. Fassnacht M, Tsagarakis S, Terzolo M, *et al.* European Society of Endocrinology clinical practice guidelines on the management of adrenal incidentalomas, in collaboration with the European network for the study of adrenal tumors. *Eur J Endocrinol*. 2023;189(1):G1-G42. https://doi.org/10.1093/ejendo/lvad066
4. Wood PJ, Barth JH, Freedman DB, Perry L, Sheridan B. Evidence for the low dose dexamethasone suppression test to screen for Cushing's syndrome–recommendations for a protocol for biochemistry laboratories. *Ann Clin Biochem*. 1997;34(Pt 3):222-229. https://doi.org/10.1177/000456329703400302
5. Blethen SL, Chasalow FI. Overnight dexamethasone suppression test: normal responses and the diagnosis of Cushing's syndrome. *Steroids*. 1989;54(2):185-193. https://doi.org/10.1016/0039-128X(89)90093-7
6. El-Farhan N, Rees DA, Evans C. Measuring cortisol in serum, urine and saliva—are our assays good enough? *Ann Clin Biochem*. 2017;54(3):308-322. https://doi.org/10.1177/0004563216687335
7. Hawley JM, Owen LJ, Lockhart SJ, *et al.* Serum cortisol: an up-to-date assessment of routine assay performance. *Clin Chem*. 2016;62(9):1220-1229. https://doi.org/10.1373/clinchem.2016.255034
8. Gant Kanegusuku A, Araque KA, Nguyen H, Wei B, Hosseini S, Soldin SJ. The effect of specific binding proteins on immunoassay measurements of total and free thyroid hormones and cortisol. *Ther Adv Endocrinol Metab*. 2021;12:2042018821989240. https://doi.org/10.1177/2042018821989240
9. Fanelli F, Cantù M, Temchenko A, *et al.* Report from the HarmoSter study: impact of calibration on comparability of LC-MS/MS measurement of circulating cortisol, 17OH-progesterone and aldosterone. *Clin Chem Lab Med*. 2022;60(5):726–739. https://doi.org/10.1515/cclm-2021-1028
10. Atkins JS, Hawley JM, Owen LJ, Clayton J, Scargill J, Keevil BG. Serum cortisol assay performance following the 1 mg overnight dexamethasone suppression test. *Ann Clin Biochem*. 2023;60(6):386-395. https://doi.org/10.1177/00045632231179560
11. Fanelli F, Belluomo I, Di Lallo VD, *et al.* Serum steroid profiling by isotopic dilution-liquid chromatography-mass spectrometry: comparison with current immunoassays and reference intervals in healthy adults. *Steroids*. 2011;76(3):244-253. https://doi.org/10.1016/J.STEROIDS.2010.11.005
12. The European Federation of Clinical Chemistry and Laboratory Medicine database. Accessed December 1, 2024. https://biologicalvariation.eu/meta_calculations
13. Oosterhuis WP, Bayat H, Armbruster D, *et al.* The use of error and uncertainty methods in the medical laboratory. *Clin Chem Lab Med*. 2018;56(2):209–219. https://doi.org/10.1515/cclm-2017-0341
14. Briegel J, Sprung CL, Annane D, *et al.* Multicenter comparison of cortisol as measured by different methods in samples of patients with septic shock. *Intensive Care Med*. 2009;35(12):2151-2156. https://doi.org/10.1007/S00134-009-1627-9
15. Monaghan PJ, Owen LJ, Trainer PJ, Brabant G, Keevil BG, Darby D. Comparison of serum cortisol measurement by immunoassay and liquid chromatography-tandem mass spectrometry in patients receiving the 11β-hydroxylase inhibitor metyrapone. *Ann Clin Biochem*. 2011;48(Pt 5):441-446. https://doi.org/10.1258/ACB.2011.011014
16. El-Farhan N, Pickett A, Ducroq D, *et al.* Method-specific serum cortisol responses to the adrenocorticotrophin test: comparison of gas chromatography-mass spectrometry and five automated immunoassays. *Clin Endocrinol (Oxf)*. 2013;78(5):673-680. https://doi.org/10.1111/cen.12039
17. Grassi G, Morelli V, Ceriotti F, *et al.* Minding the gap between cortisol levels measured with second-generation assays and current diagnostic thresholds for the diagnosis of adrenal insufficiency: a single-center experience. *Hormones (Athens)*. 2020;19(3):425-431. https://doi.org/10.1007/s42000-020-00185-y
18. Javorsky BR, Raff H, Carroll TB, *et al.* New cutoffs for the biochemical diagnosis of adrenal insufficiency after ACTH stimulation using specific cortisol assays. *J Endocr Soc*. 2021;5(4):bvab022. https://doi.org/10.1210/jendso/bvab022
19. Dodd AJ, Ducroq DH, Neale SM, *et al.* The effect of serum matrix and gender on cortisol measurement by commonly used immunoassays. *Ann Clin Biochem*. 2014;51(3):379-385. https://doi.org/10.1177/0004563213514567
20. Raverot V, Richet C, Morel Y, Raverot G, Borson-Chazot F. Révision des seuils diagnostics à utiliser avec le nouveau réactif de dosage du cortisol plasmatique de roche diagnostics. *Ann Endocrinol (Paris)*. 2016;77(5):620-622. https://doi.org/10.1016/j.ando.2016.05.002
21. Kline GA, Buse J, Krause RD. Clinical implications for biochemical diagnostic thresholds of adrenal sufficiency using a highly specific cortisol immunoassay. *Clin Biochem*. 2017;50(9):475-480. https://doi.org/10.1016/J.CLINBIOCHEM.2017.02.008
22. Di Dalmazi G, Vicennati V, Garelli S, *et al.* Cardiovascular events and mortality in patients with adrenal incidentalomas that are either non-secreting or associated with intermediate phenotype or subclinical Cushing's syndrome: a 15-year retrospective study. *Lancet Diabetes Endocrinol*. 2014;2(5):396-405. https://doi.org/10.1016/S2213-8587(13)70211-0
23. Debono M, Bradburn M, Bull M, Harrison B, Ross RJ, Newell-Price J. Cortisol as a marker for increased mortality in patients with incidental adrenocortical adenomas. *J Clin Endocrinol Metab*. 2014;99(12):4462-4470. https://doi.org/10.1210/JC.2014-3007

24. Puvaneswaralingam S, Kjellbom A, Lindgren O, Löndahl M, Olsen H. ACTH following overnight dexamethasone suppression can be used in the verification of autonomous cortisol secretion in patients with adrenal incidentalomas. *Clin Endocrinol (Oxf)*. 2021;94(2):168-175. https://doi.org/10.1111/CEN.14357
25. Sconfienza E, Tetti M, Forestiero V, Veglio F, Mulatero P, Monticone S. Prevalence of functioning adrenal incidentalomas: a systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2023;108(7):1813-1823. https://doi.org/10.1210/CLINEM/DGAD044
26. Vogeser M, Kratzsch J, Ju Bae Y, *et al*. Multicenter performance evaluation of a second generation cortisol assay. *Clin Chem Lab Med*. 2017;55(6):826-835. https://doi.org/10.1515/cclm-2016-0400
27. Nenke MA, Lewis JG, Rankin W, Torpy DJ. Evidence of reduced CBG cleavage in abdominal obesity: a potential factor in development of the metabolic syndrome. *Horm Metab Res*. 2016;48(8):523-528. https://doi.org/10.1055/s-0042-108728
28. Vogg N, Kurlbaum M, Deutschbein T, Gräsl B, Fassnacht M, Kroiss M. Method-specific cortisol and dexamethasone thresholds increase clinical specificity of the dexamethasone suppression test for cushing syndrome. *Clin Chem*. 2021;67(7):998-1007. https://doi.org/10.1093/clinchem/hvab056